冗袅データ配億ユニットを通じてデータをカスケードにする方法および装置

3 (19) 日本国格許庁 (JP)

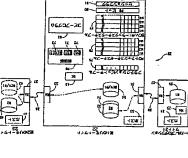
3 扆 ধ

P.1 G.0 6F 12/00 5.3 1D 5.B 0.1 8 5.3 5.R 5.B 0.8 2 12/16 3.1 0.9 5.B 0.8 3 13/00 3.0 1.P	等正对决 有 子撰審查謝求 未請求(全 43 頁)	(71) 出観人 イーエムシー コーボレイション アメリカ合衆国 マサチューセッツ州 01748 ホブキントン サウス ストリート 171 (72)発明者 ガーニュ マチュー アメリカ合衆国 マサチューセッツ州 0214 ポストン ジョイ ストリート 43 アバートメント 5 43 アバートメント 5 7メリカ合衆国 マサチューセッツ州 アメリカ合衆国 マサチューセッツ州
MBMR245 F 1 631 G0 535 310 301	· 排	(21) 出頭番骨 特麗2000-600177(P2000-600177) (71) (86) (22) 出版相 平成11年10月12日(1999, 10.12) (85) 翻改法指出日 平成12年10月17日(2000, 10.17) (86) 国際以間審身 PCT/US 9 9 / 2 3 9 3 9 (72) 国际公開番号 WCO 0 / 4 9 5 0 0 (37) 國先組主張書号 WCO 0 / 4 9 5 0 0 (37) 國先組主張書号 0 9 / 2 5 1, 8 1 2 (32) 優先組主張書号 0 9 / 2 5 1, 8 1 2 (32) 優先組主張書号 EP(AT, BE, CH, CY, DE, DK, ES, F1, FR, GB, GR, IE, I
(51) Int.Cl.? G 0 6 F 12/00 12/16 13/00		(81) 出頭部母 (86) (22) 出版用 (85) 翻訳文指出目 (85) 國歌文指出目 (87) 國歌文開書母 (87) 國歌大聞主張書母 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相 (32) 優先相

冗長ゲータ記憶ユニットを通じてデータをカスケードにする方法および装置 [発別の名称] (24)

タ変更装置からリモートデータ受信サイト(23)ヘデ --タを転送するためのデータ記憶影像(2.2)。 このデ **-夕配億股億は、データ変更装置によって生成されたデ** の変更をトラック年に記録する。第2のデータ記憶装置 のデータ記憶装置のミラーになる。 第2の動作モードの **問中に、第2のデータ記憶装置は、ミラーとして働くこ** とをやめ、データ受信サイトへのデータの転送のための ノースになる。変更された情報のみ、すなわち変更され た特定のトラックが、第2の動作モード中の逆続した動 トからのコマンドは、第1と第2の動作モードの間の転 --タ中のそれぞれの変更を記録するための第1のデータ 記憶装置を含む。レジスタセット (75) は、それぞれ (34) は、第1及び第2の動作モードを有する。第1 作の間中に概形される。ローカルのプロダクションサイ プロダクションデータ処理サイト (21) のようなデー の動作モードの間中に、第2のデータ記憶装置は、第1 [五変] 23

送を開始する。



最終買に扱く 平成14年12月10日(2002, 12, 10) **特表2002-542526** (P2002-542526A) (11)特許出國公教每号 (2) B (2) ¥ tr E (43)公表日 宇斯 田田 田田 辍 4 盐 称

* [45.1]

冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

26

ഗ

4 2

വ 1

特表2002

[特許請求の範囲]

データ変更手段とデータ受信 データ処理システムにおける、 手段の間に介在するデータ記憶設備であって、 [請求項1]

データを前記データ変更手段から受信するための、前記データ変更手段との 接続のための第1のデータ記憶装置と、

- B. 第2のデータ記憶装置と、
- 前記第1のデータ記憶装置中で前記データ変更手段が行う変更を識別するた めの変更記録手段と、 ပ်
- D. 前記変更記録手段中に記録された変更に従って、第1の動作モードの間に前 記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信する、 のような第1の動作モードを確立するための第1の動作手段と、
- 第2の動作モードが確立されるときに前記変更記録手段中に記録される変更 めの前記第2の動作モードを確立するための第2の動作手段と、を有することを に従って、前記第2のデータ記憶装置から前記受信手段にデータをコピーするた 特徴とするデータ記憶設備。

それぞれのデータブロックについての別の情 前記別の情報を受信するため データプロックとして起こ の第1のテープルを含み、前記第2の動作手段は、所定の情報の内容を前記第1 のデータ記憶装置から前記第1のテープルに転送することを特徴とする謂求項: 前記データ記憶設備内の転送は、 報のための記憶装置を含み、前記変更記録手段は、 り、前記第1のデータ記憶装置は、 に記載のデータ記憶設備。 [請求項2]

前記変更記録手段に転送することを特徴とする請求項1に記載のデータ記憶設備 り、前記第1のデータ記憶装置の部分は、それぞれのデータブロックについての 情報を含み、前記第2の動作手段は、前記第1のデータ記憶装置の部分の内容を 【静求項3】 前記データ記憶設備内の転送は、データブロックとして起こ

前記データ記憶設備は、前記第1及び第2の動作手段を相互 に排他的に使用可能にする前記変更手段から受信する所定のデータに応答する制 御手段を更に含むことを特徴とする請求項3に記載のデータ記憶設備。 [請求項4]

前記データ記憶設備は、カスケードセッションの間中に制御 能にする、そのようなカスケードセッションを実行するための前記変更手段から 手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可 の、反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に 含むことを特徴とする請求項3に記載のデータ記憶設備。 請求項5】

データを転送するためのコピー手段を更に含み、前記変更記 こおける最初の反復の間中に、前記第2のテーブル中の全てのデータを第1の状 それによって前記コピー手段が前記第2のデータ記憶装置中の全 ての前記データを前記受信手段に転送することを特徴とする請求項5に記載のデ 録手段は、第1、第2及び第3のテーブルを含み、前記第1の動作手段は、所定 の位置から前記第1のテーブルヘデータを転送し、かつ、カスケードセッション [請求項6] 態にセットし、 ータ記憶設備。

次の動作における前記第1の動作手段は、前記第1及び第2 のテーブル中の対応するデータを変更し、次の動作における前記第2の動作手段 は、前記第1のテーブルから前記第3のテーブルに対応するデータを転送するこ とを特徴とする請求項6に記載のデータ記憶設備。 [請求項7]

田田田 【請求項8】 ホスト及びプロダクション記憶設備を有するプロダクション サイトを含み、かつ、リモート記憶設備、前記プロダクションサイトを相互接続 するための記憶設備、及び前記ホストが複数のカスケードコマンドを発行するこ とができるような前記リモートサイトを含むデータネットワークにおける、 記憶設備であって、

A. トラック毎にデータを前記プロダクション設備から受信するための、前記プ ロダクション設備との接続のための第1のデータ記憶装置と、

- 第2のデータ記憶装置と、 . В
- C. 前記プロダクション設備が変更する前記第1のデータ記憶装置中のそれぞれ のトラックを識別するための複数のトラックステータステーブルと
- D. 第1のカスケードコマンドに応答して、前記トラックステータステーブル中 に記録された変更に従って、第1の動作モードの間中に前記第2のデータ記憶装 置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モ

冗長データ記憶ユニットを通じてデータをカス ケードにする方法および装置

9 2

Ŋ

8

4 S Į

~ 0

特表20

ドを確立するための第1の動作手段と、

E. 第2のカスケードコマンドに応答して、第2の動作モードが確立されるとき に前記トラックステータステーブル中に記録される変更に従って、前記第2のデ の動作モードを確立するための第2の動作手段と、を有することを特徴とする中 一夕記憶装置から前記リモート記憶設備にデータをコピーするための、前記第2 間記憶設備。 前記第1のデータ記憶装置は、それぞれのトラックについて の別の情報のための記憶装置を含み、前記トラックステータステーブルは、前記 の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送するこ 別の情報を受信するための第1のテーブルを含み、前記第2の動作手段は、 とを特徴とする請求項8に記載の中間記憶設備。 [請求項9]

ラックについての情報を含み、前記第2の動作手段は、前記第1のデータ記憶装 置の部分の内容を前記トラックステータステーブルに転送することを特徴とする 前記第1のデータ記憶装置の部分は、それぞれのデータ 請求項8に記載のデータ記憶設備。 [請求項10]

に排他的に使用可能にする前記プロダクション設備から受信する所定のデータに 【請求項11】 前記中間記憶設備は、前記第1及び第2の動作手段を相互 応答する制御手段を更に含むことを特徴とする請求項10に記載のデータ記憶設

前記中間記憶設備は、カスケードセッションの間中に制御 手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可 ン設備からの、反復して生成される第1及び第2のコマンドに応答する前記制御 そのようなカスケードセッションを実行するための前記プロダクシ 手段を更に含むことを特徴とする請求項10に記載のデータ記憶設備。 [請求項12] 能にする、

を転送し、かつカスケードセッションにおける最初の反復の間中に、前記第2の データを転送するためのコピー手段を更に含み、前記トラ ックステータステーブルは、第1、第2及び第3のテーブルを含み、前記第1の 動作手段は、前記第1のデータ記憶装置の部分から前記第1のテープルへデータ テーブル中の全てのデータを第1の状態にセットし、それによって前記コピ [請求項13]

段が前記第2のデータ記憶装置中の全ての前記データを前記受信手段に転送する ことを特徴とする請求項12に記載のデータ記憶設備。 、請求項14】 次の動作における前記第1の動作手段は、前記第1及び第 2のテーブル中の対応するデータを変更し、次の動作における前記第2の動作手 段は、前記第1のテーブルから前記第3のテーブルに対応するデータを転送する ことを特徴とする請求項13に記載のデータ記憶設備。 前記中間記憶設備は、前記リモートサイトから地理的に遠 ための手段を更に有することを特徴とする請求項10に記載のデータ記憶設備。 隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立す [請求項15]

【請求項16】 前記中間記憶設備は、前記プロダクション設備から地理的 **確立するための手段を更に有することを特徴とする請求項10に記載のデータ記** に遠隔であり、前記プロダクション設備と前記中間記憶設備の間の通信リンク

前記中間記憶設備は、前記リモートサイトから地理的に遠 ための手段を更に有することを特徴とする請求項16に記載のデータ記憶設備。 脳であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立す [請求項17]

前記中間記憶設備は、前記リモートサイトから地理的に遠 隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立する ための手段を更に有し、前記通信リンクのそれぞれは、異なる特性のバンド幅を 有することを特徴とする請求項16に記載のデータ記憶設備。 (請求項18]

【請求項19】 前記中間記憶設備は、前記リモートサイトから地理的に遠 モートサイトへの前記通信リンクより広い特性のバンド幅を有することを特徴と 隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立する ための手段を更に有し、前記プロダクション設備への前記通信リンクは、 する請求項16に記載のデータ記憶設備。 データ変更設備とデータ受信設備の間でデータを転送す ための方法であって、 [請求項20]

A. 変更されたデータを受信するための第1のデータ記憶装置を確立するステッ

特表2002 冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

9

~

S

-542

- 第2のデータ記憶装置を確立するステップと
- 受信された変更されたデータによって作り出される前記第1のデータ記憶装 置中の変更を識別するステップと、
- D. 識別された前記変更に従って、第1の動作モードの間に前記第2のデータ記 そのような第1の動 **飯装置が前記第1のデータ記憶装置からデータを受信する、** 作モードを確立するステップと、
- 第2の動作モードが確立されるときに識別される変更に従って、前記第2の データ記憶装置から前記受信設備にデータをコピーするための第2の動作モード を確立するステップと、を有することを特徴とする方法。
- 【静求項21】 前記データ記憶設備内の転送は、データブロックとして起 こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の 情報のための記憶装置を含み、前記変更を識別するステップは、前記別の情報を 受信するための第1のテーブルを確立するステップ、及び前記第2の動作モード が確立しているときに所定の情報の内容を前記第1のデータ記憶装置から前記第 1のテーブルに転送することを特徴とする請求項20に記載の方法。
- 前記データ記憶設備内の転送は、データブロックとして起 の情報を含み、前記第1のデータ記憶装置の部分の内容は、最初の識別情報を提 こり、前記第1のデータ記憶装置の部分は、それぞれのデータブロックについて **供するために転送されることを特徴とする請求項20に記載の方法。** [請求項22]
- 【間求項23】 前記第1及び第2の動作モードは、前記変更設備からの所 定のデータに広答して相互に排他的に動作することを特徴とする請求項22に記 載の方法。
- 成は、カスケードセッションの間中に前記第1及び第2の動作モードが相互に排 前記変更設備からの第1及び第2のコマンドの反復する生 そのようなカスケードセッションを実行する ことを特徴とする請求項22に記載の方法。 他的に反復して使用可能にされる、 [請求項24]
- 第1、第2及び第3のテーブルを含み、第1の動作モードの開始は、所定の位置 から前記第1のテーブルヘデータを転送し、かつカスケードセッションにおける (請求項25] データは、コピーによって転送され、変更の前記識別は、

- 2 -

最初の反復の間中に、前記第2のテーブル中の全てのデータを第1の状態にセッ トし、それによって前記コピーが前記第2のデータ記憶装置中の全ての前記デー タを前記受信設備に転送することを更に特徴とする請求項24に記載の方法。

「請求項26】 次の動作における前記第1の動作モードは、前記第1及び 第2のテーブル中の対応するデータを変更し、次の動作における前記第2の動作 モードは、前記第1のテーブルから前記第3のテーブルに対応するデータを転送 することを特徴とする請求項25に記載の方法。 ホスト及びプロダクション記憶設備を有するプロダクショ ンサイトと、リモート記憶設備を有するリモートサイトの間でデータを転送する ための方法であって、前記ホストは、複数のカスケードコマンドを発行するこ。 [請求項27] がつず

A. トラック毎にデータを前記プロダクション設備から受信するための第1のデ ータ記憶装置を確立するステップと

- 第2のデータ記憶装置を確立するステップと、
- のトラックを識別するための複数のトラックステータステーブルを定義するステ 前記プロダクション設備が変更する前記第1のデータ記憶装置中のそれぞれ ップと
- D. 第1のカスケードコマンドに応答して、前記トラックステータステーブル中 置が前記第1のデータ記憶装置からデータを受信する、そのような第1の動作モ に記録された変更に従って、第1の動作モードの間中に前記第2のデータ記憶装 一ドを確立するステップと、

に前記トラックステータステーブル中に記録される変更に従って、前記第2のデ E. 第2のカスケードコマンドに応答して、第2の動作モードが確立されるとき ータ記憶装置から前記リモート記憶設備にデータをコピーするための、前記第2 の動作モードを確立するステップと、を有することを特徴とする方法。 【請求項28】 前記第1のデータ記憶装置は、それぞれのトラックについ 前記別の情報を受信するための第1のテーブルを含み、前記第2の動作手段は 所定の情報の内容を前記第1のデータ記憶装置から前記第1のテーブルに転送 ての別の情報のための記憶装置を含み、前記変更トラックステータステーブルは

することを特徴とする請求項27に記載の方法。

前記第1のデータ記憶装置の部分は、それぞれのトラック についての情報を含み、前記第2の動作手段は、前記第1のデータ記憶装置の部 分の内容を前記トラックステータステーブルに転送することを特徴とする請求項 [請求項29] 7に記載の方法。

を相互に排他的に使用可能にする前記プロダクション設備から受信する所定のデ 前記中間データ記憶設備は、前記第1及び第2の動作手段 ータに応答する制御手段を更に含むことを特徴とする請求項29に記載の方法。 [請求項30]

使用可能にする、そのようなカスケードセッションを実行するための前記プロダ クション設備からの、反復して生成される第1及び第2のコマンドに応答する前 前記中間データ記憶設備は、カスケードセッションの間中 に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に 記制御手段を更に含むことを特徴とする請求項29に記載の方法。 [請求項31]

2及び第3のテーブルを有する前記トラックステータステーブルを確立し、前記 態に前記第2のテーブル中のすべてのデータをセットし、それによって前記コピ 一手段が前記第2のデータ記憶装置中の全ての前記データを前記受信手段に転送 第1、第 第1のデータ記憶装置の部分から前記第1のテーブルヘデータを転送し、かつカ スケードセッションにおける最初の反復の間中に、第1の動作モードの第1の状 【請求項32】 データ転送を実行するためにデータをコピーし、 するステップを更に含むことを特徴とする請求項31に記載の方法。

次の反復における前記第1の動作モードは、前記第1及び 次の反復における前記第2の動作 モードの間に、前記第1のテーブルから前記第3のテーブルに対応するデー 転送されることを特徴とする請求項32に記載の方法。 第2のテーブル中の対応するデータを変更し、 [請求項33]

前記中間記憶設備は、前記リモートサイトから地理的に遠 隔であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立す ステップを更に有することを特徴とする請求項29に記載の方法。 [請求項34]

(請求項35) 前記中間記憶設備は、前記プロダクション設備から地理的 に遠隔であり、前記プロダクション設備と前記中間記憶設備の間の通信リンク

淄立するステップを更に有することを特徴とする請求項29に記載の方法。

前記中間記憶設備は、前記リモートサイトから地理的に遠 脳であり、前記中間記憶設備と前記リモートサイトの間の通信リンクを確立する ステップを更に有することを特徴とする請求項35に記載の方法。 [請求項36]

前記通信リンクを確立する前記ステップは、異なる特性の バンド幅を有する通信リンクを確立することを特徴とする請求項35に記載の方 [請求項37] 炭

前記リモート サイトへの前記通信リンクより広い特性のバンド幅を有する前記プロダクション 設備への通信リンクを確立することを特徴とする請求項35に記載の方法 前記通信リンクを確立する前記ステップは、 [請求項38]

[発明の詳細な説明]

[0001]

冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

(関連出願とのクロス・リファレンス)

たリモートデータ設備の独立動作の方法及び装置(Method and Apparatus for In 出願がこの発明と同じ譲受人に譲渡された、1996年5月31日に出願され dependent Operation of a Remote Data Facility)のための米国特許第08/6 56,035号。

[0002]

and Apparatus for Independent and Simultaneous Access to a Common Data S 出願がこの発明と同じ譲受人に譲渡された、1997年4月25日に出願され た共通データセットへの独立かつ同時のアクセスのための方法及び装置 (Method et)のための米国特許第08/656,035号。

[0003]

(技術分野)

より詳しくは、複数の冗長データ記憶ユニット(redundant data storage unit)を通じてデータをカスケードにする(cascade)ことを可能にする方法及び装置に この発明は、広くデータ処理ネットワーク中の冗長データ記憶デバイスに関し

[0004]

(背景技術)

データの冗長性によるデータの健全性の維持は、非常に重要な問題となってい 656,035号などに記載された他の形態では、冗長性は、遠隔の位置にある ローカルなシステムを複製することによって達成される。遠隔の場所を使用する ことによって、1つの場所での自然災害などによってデータが失われることが防 ディスク又は他の記憶 装置の故障から保護する。1996年5月31日に出願された米国特許第08/ る。データの冗長性は、いくつかの形態及びバリエーションを有する。 イトでは、ミラリング(mirroring)又はRAIDの冗長性が、 止される。

[0005]

ဖ

52

先の米国特許第08/656,035号によると、すべてのデータ処理活動は

データ記憶設備中に記憶されたデータを処理するためのホストシステムを含む

地理的 に遠隔な又はバックアップのサイトは、復元(restoration)又は回復(recovery) 「ローカル」の又は「プロダクション(production)」のサイトで起こる。

データセットを維持するための「冗長な」設備として、データ記憶設備を含む。

このシステムでは、プロダクションサイトでホストがプロダクションサイトデー

ト記憶設備にデータを自動的に書き込む。これらのアプリケーションの、ほとん 夕記憶設備にデータを書き込む度に、プロダクションデータ記憶設備は、リモー

どではないとしても多くでは、リモートサイトのデータ記憶設備にデータを書き 込むことは、バックアップ手順がプロダクション設備の運転に影響を与えないよ うに、広いバンド幅の通信リンクを通した転送を必要とする。T3及びESCONライ

それらの費用にもかかわらず、通信リンクとして、典型的に好まれている ばべ

[0000]

953号は、プロダクションサイトで典型的に使用 ット、例えば、プロダクションサイトのデータ記憶設備上の「論理ボリューム(1 ボリュームと同期する。SPLITコマンドは、その後、BCVボリュームをプロダクシ Vボリューム(BCV volume)」として示される。ESTABLISHコマンドは、BCVボリュ されるBCVデバイスを含む他の概念を開示する。このアプローチでは、データセ ームをプロダクションボリュームに接続し、BCVボリュームは、プロダクション ogical volume)」、すなわち「プロダクションボリューム(production volume) ョンボリュームから分離することができ、BCVボリューム上に記録されたデータ 」は、通常の方法で動作する。この同じサイトでの他の論理ボリュームは、 を他のアプリケーションで利用できるようにすることができる。 米国特許第08/842,

[0000]

ることができる。結果として、BCVボリュームがプロダクションボリュームと再 接続されるときはいつでも、BCVボリュームとプロダクションボリュームの両方 この他のアプリケーションは、BCVポリューム上に記憶されたデータを変更す の変化を表わすデータをBCVポリュームに転送する必要がある。他の選択肢は、

特表2002-冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

9

2

Ŋ

542

別のESTABLISHコマンドを出すことによって、BCVボリュームをプロダクションボ BCVボリューム上 プロダクションボリュームからのすべてのデータが、BCVボリュームに転送され のすべてのデータを置換するであろう。変化が起こったかどうかにかかわらす、 このコマンドは、 リュームに再接続することである。しかし、 なければならない。

[0008]

캒 反復することは、過度の通信コストを含む。プロダクションサイトと第2のリモ 容されてきた。特定の重要な用途では、第1の2つのサイトから遠隔の第3のサ イトで更に他の復元コピーを記憶させることが今でも望まれ、またいくつかの状 兄では、それは義務的である。これによって、すべてプロダクションサイトの動 トにデータをコピーするためのアプローチが必要となる。プロダクションサイト から単一のリモートサイトにデータをコピーするための前述のアプローチを単に ートサイトの間、あるいは第1と第2のリモートサイトの間のいずれかで、第2 の高速通信リンクが必要となるであろう。必要なものは、よりコストの低い通信 リンクを通じた、種々の遠隔に位置するサイトで存在するデータに重大な相違が 全くない、すべてのデータ記憶設備が同期するか又はほぼ同期するような、冗長 作に透過的に(transparently)、プロダクションサイトから両方のリモートサイ 過去では、プロダクションサイトでのデータ記憶設備の位置を示すことは、 なデータのコピーを確立する方法である。

[0000]

(発明の開示)

ができる冗長なデータ記憶設備を確立するための方法及び装置を提供することで 従って、本発明の目的は、よりコストの低い通信リンクを通じて通信すること

$[0\ 0\ 1\ 0]$

本発明の他の目的は、地理的に遠隔なサイトで、冗長なディスク記憶を実現す 経済的に提供することである。 るための方法及び装置を、

[0011]

本発明の更に他の目的は、経済的な通信リンクを使用して、2つ以上の地理的

に遠隔なサイトに、冗長なデータ記憶を提供するための方法及び装置を提供する ことである。

[0012]

2つ以上の地理的に遠隔なサイトに、冗長なデータ記憶を提供するための方法 本発明のまた更に他の目的は、経済的な低いパンド幅の通信リンクを使用して 及び装置を提供することである。

[0013]

プロダクションサイドでの動作に透過的に達成される冗長性で、2,つ以上の冗長 本発明のまた更に他の目的は、プロダクションサイトで処理されるデータが、 な記憶サイトに記憶されるデータ処理ネットワークを提供することである。

[0014]

本発明によると、プロダクション設備のような、データを変更するデータ処理 中で記録された変更に従って第1のデータ記憶装置からデータを受信する、その 設備から、リモートデータ記憶サイトは、データ記憶設備から遠隔の他のサイト トからデータを受信する。データ変更レコーダ(data change recorder)は、デー タ処理設備が第1のデータ記憶装置中で行った変更を識別する。第1の動作コン ような第1の動作モードを確立する。この動作が完了した後、第2の動作コント ロールは、前記データ変更レコーダ中に記録された変更に従って、第2のデータ 記憶装置から他のリモートサイトにデータをコピーするための第2の動作モード 第1及び第2のデー タ記憶装置(data store)を含む。第1のデータ記憶装置は、プロダクションサイ トロール(operating control)は、その動作モードの間に前記変更レコーダ手段 にデータを転送する。リモートサイトのデータ記憶設備は、 を確立する。

[0015]

トの間で転送される。プロダクションサイトは、ホスト及びプロダクション記憶 設備を含み、リモートサイト、リモート記憶設備は、トラック毎に(track-by-tr ack basis)プロダクション設備からデータを受け取ることを特徴とする第1のデ ータ記憶装置、及び第2のデータ記憶装置を含む。ホストは、プロダクション設 本発明の他の態様によると、データは、プロダクションサイトとリモートサイ

特表2002-542 冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

9

52

クステータステーブル中に記録された変更に従って、第1のデータ記憶装置から データを受信する。第2の動作モードの間では、第2のデータ記憶装置からのデ **一夕は、第2のモードが確立されるときに、トラックステータステーブル中に記** ンドを発行することができる。更に、ホストは、第1及び第2の動作モードを確 立することができる。第1の動作モードでは、第2のデータ記憶装置は、トラッ 複数のトラックステータステーブルの画定を可能にする、複数のカスケードコマ **備が変更する第1のデータ記憶装置中のそれぞれのトラックを識別するための** 碌される変更に従って、リモート記憶設備にコピーされる。

$[0\ 0\ 1\ 6]$

(図面の簡単な説明)

添付の請求項は、本発明の主題を、特に指し示し、明確にクレームする。本発 そのような添付の図面と共に、以下の詳細な説明を読むことから、より完全に明 明の種々の目的、利点及び新しい特徴は、同様の参照番号は同様の部品を示す、 らかになるであろう。

[0017]

(本発明を実施するためのベストモード)

(一般的動作)

を示す。第1及び第2のリモートサイト22及び23は、典型的には、ローカル のプロダクションサイト21から、及びお互いから、地理的に離れているであろ 図1は、ローカル又はプロダクションサイト21、第1のリモートサイト22 及び第2のリモートサイト23に装置が設置されたデータ処理ネットワーク20 ローカルのプロダクションサイト21 一緒に置くことができるであろう。 及び第1のリモートサイト22の装置は、 う。しかし、明らかになっていくように、

[0018]

冗長性の第1のレベルは、ローカルのプロダクションサイト21と第1のリモ ンメモリを含むホスト24は、種々のプログラムで動作する。周期的に、ホスト 2.4は、ホストアダプタ25を通じてディスク記憶デバイスへの転送を実行する 20中で達成される。知られているように、1つ以上の中央プロセシサ及びメイ ートサイト22の間の相互作用(interaction)を通じてデータ処理ネットワーク

れらは、ファイル又は他のデータブロックを含む。本発明の譲受人によって製造 ラックを含む論理ボリュームであり、転送はトラック毎に実行される。図1では の論理ボリュームを表わす。明らかになっていくように、ホスト24は、それが そのようなプロダクション記憶設備からホストアダプタ25を通じて必要とする であろう。このディスク記憶デバイスは、関連する情報を記憶するための別個の 、R1論理ボリューム26は、通常そのようなディスク記憶設備中に含まれる多く される装置の状況では、典型的な記憶セクションは、多くの隣接したディスク セクションに編成された多くの物理ディスクドライブを有することができる。 どのような情報も検索する。そのようなシステムは、技術上周知である。 1つのそのようなシステムを開示す 許第08/656,035号は、

ミラリングしたり、又は種々のRAID構成によってホスト24による割込 み動作から特定の物理ディスクドライブの機能不全を防止する、そのような種々 のRAID構成のいずれか1つで、論理ボリュームのアレイによって構成したりする 示されてはいないが、ローカルのプロダクションサイト21は、R1論理ボリュ ことができる。明らかになっていくように、ローカルのプロダクションサイト2 1 で実施されるどのような冗長性の構成にも関わらず、本発明を適用することが ーム26のための冗長性を提供することができる。例えば、R1ボリュームを、 際には、 るまる。

[0020]

を通じて、第1のリモートサイト22中の対応するリモートアダプタ31に接続 24がR1論理ボリューム26にデータを書き込むごとに、R2論理ボリューム32 544,347号に記載されているように、第1のリモートサイト22でN1論理 ボリューム26中のデータをミラリングすることによって実施することができる ローカルの この構成では、ホスト プロダクションサイト21は、ESCON又はT3通信ラインのような高速通信リンク への転送のためにそのデータを高速通信リンクを通じて第1のリモート記憶サイ 技術上知られ、本発明と同じ護受人に譲渡されたヤナイ他への米国特許第5, 他の冗長性の構成が開示される。この構成においてより具体的には、 するリモートアダプタ(remote adapter, RA)27を含む。

特表2002-542 冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

ၑ 2

ည

なわち伝達される。もし、自然災害がローカルのプロダクションサイト21を襲 更に、もしホスト33が第1のリモートサイト22中に位置するなら、すべての ム26に関して遠隔のミラーとしての役割をするR2論理ボリューム32に反映す 動作は、検出されないデータの損失が全くないまま、連続した動作のために第1 ったなら、データは、リモートサイトのR2論理ボリュームで簡単に利用できる。 ト22中のリモートアダプタ31に転送することによって、リモートアダプタ 7は応答する。このように、それぞれのWRITE(書込み)動作は、R1論理ポリュー のリモートサイトにシフトすることができる。

[0021]

前述のように、米国特許出願第842,953号は、データ記憶設備に関する BCV/R1論理ボリューム34は、第1のリモートサイト22中に含まれる。それ は、第1のリモートサイト22内の、好適にはR2論理ボリューム32を含む物理 ディスクドライブと異なる物理ディスクドライブ上に、任意の専用の論理ポリュ BCV論理ボリュームを含むデータ処理ネットワークを開示する。本発明によると 一ムを含むことができる。

[0022]

リューム32と同期する。BCV/R1論理ボリューム34がリモートアダプタ35に サイト23中のリモートアダプタ37に転送されるであろう。本発明の1つの実 また第1のリモ 接続している第2の動作モードでは、データは、R2論理ボリューム40又は他の データレシーバへの転送のため、他の通信リンク36を通じて、第2のリモート この発明によると、また後でより詳細に説明するように、BCV/R1論理ボリュー ム34を、R2論理ボリューム32又はリモートアダプタ35のいずれかに接続す ることができる。第1の動作モードでは、BCV論理ボリューム34は、R2論理ボ h2] 胀 一トサイト22にホスト33が存在する必要性をなくす。このようにして、 のリモートサイト23は、もし自然災害がローカルのプロダクションサイ で起きると、復元サイト又はデータで動作するための第2のサイトになる。 ホスト41を含み、 第2のリモートサイト23は、 施形態では、

[0023]

第2のリモートサイトは、オプションのBCV/RI論理ボリューム42を含むよう

論理ボリュームを含ませることによって、第3のリモートサイトへの本発明に従 第2のリモートサイト23にこの って実施される機能の複製(replication)ができる。 に示される。後でより完全に説明するように、

[0024]

一般的な条件では、本発明に従って構成されるリモートサイトは、データ記憶 れは、データを変更することができるローカルのプロダクションサイト21への ಜ の間にR2論理ボリューム32がリモートアダプタ31からデータを受信し、それ N/RI論理ポリューム34は、第2のデータ記憶装置を構成する。第1のコマンド に応答して、第1のリモートサイト22中のデータ記憶設備は、その動作モード によってRI論理ボリューム26中のデータになされた変更に応答する、そのよう な第1の動作モードで動作する。典型的には、これは、同期的に実施され、その ため高速通信リンク30が必要である。この動作モードでは、BCV/R1論理ポリュ 接続のために、R2論理ポリューム32の形態の第1のデータ記憶装置を含む。 設備を構成する第1のリモートサイト22の基本的な構造を有するであろう。 ーム34は、それのBCV又は第1の動作モードで動作していると考えられる。

业 第1の動作モードで同 モートアダプタ35、通信リンク36及びリモートアダプタ37を通じて、R2論 埋ボリューム40にデータを転送する。第1から第2の動作モードへのシフトの タイミングは、システムのオペレータによって決定されるであろう。しかし、第 典型的には、第2の動作モ の動作モードでは、第1のリモートサイトは、BCV/RI論理ボリューム34からリ 一般的には、R2論理ボリュ 账 期が達成されると、第2の又はR1の動作モードにシフトすることができる。 ードへのシフトの間の間隔は、R2論理ボリューム32中の括動によって、 本発明によると、BCV/R1論理ポリューム34は、一旦、 ーム40へのデータ転送が完了した後に実施される。 2の動作モードから第1の動作モードへのシフトは、 間又は更には日の期間になるであろう。

[0026]

更に本発明によると、BCV/RI論理ポリューム34が、それの第2の動作モード にシフトするたびに、第1の動作モードの間に変更されたデータトラックのみが

特表2002 冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

9

5 2

4 2

ر ا

のトラックへの反復した変更をするなら、BCV/RI論理ポリューム34からR2論理 リンク36を、低いバンド幅の電話線又はインターネットを通じて転送すること ポリューム40への1つの転送のみが起こるであろう。そのような減少によって R2論理ボリューム40に転送される。もしローカルのプロダクションサイト 1が、第2の動作モードへの連続したシフトの間にR2論理ポリューム32の単 通信リンク36へのバンド幅の要件を減少させることができる。例えば、 ができるようなレベルに減少させることができるであろう。

[0027]

このように、本発明に従って、81論理ボリューム26になされたデータの変更 は、R2論理ボリューム32中で、そしてBCV/R1論理ボリューム34を通じてR2論 **埋ポリューム40へと複数される。順番にデータを転送するこの処理は、カスケ** ドとして特徴付けられ、ローカル又はプロダクションサイト21でホスト24 によって処理されるカスケードコマンドセットによって制御される。

[0028]

(カスケードコマンド処理)

ト21で処理されているアプリケーションプログラムは、第1のリモートサイト 22で実行されているカスケード動作を実行する。知られているように、図1の ホスト24は、カスケードコマンドセット中のコマンドを含む種々のコマンドを 作り出すであろう。これらのコマンドは、ホストアダプタ25を含む任意の数の より具体的には、システムのオペレータ、又はローカルのプロダクションサイ デバイス又はアドレスに命令することができる。

[0029]

ここで図1及び2を参照すると、ホストアダプタ25がコマンドを受信すると き、ステップ44は、そのコマンドをテストしてデコードするために普通の処理 を使用する。そのテストは、正しいコマンドが受信されたかどうかを確定するた し、それがカスケードコマンドでなければ、ホストアダプタ25は、対応する ステップ45は、そのコマンドが、カスケードコマンドセット中のコマンドの1 しであるかどうか、すなわち、カスケードコマンドであるかどうかを確定する。 めに、文法、コンテキスト(context)及び他の関連するパラメータを分析す

特表2002· 冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置 機能を実施するために手順46を使用する。有効なカスケードコマンドを受信す ると、ホストアダプタ25は、そのコマンドをリモートアダプタ27に転送す ために、ステップ47を使用する。

[0800]

1は、それ以上分析することなく、そのコマンドを通信リンク30を通じてリモ ートアダプタ31に、単に転送する。リモートアダプタ31は、同様の処理を実 及びそのコマンドをデバイスコントローラ43に転送するためにステップ53 **行する。すなわち、リモートアダプタ31は、ステップ52でコマンドを受信し** リモートアダプタ27がステップ50でコマンドを受信したとき、ステップ を使用する。

[0.031]

ンドを受信した後、デバイスコントローラ34は、ステップ55でコマンドをデ デバイスコントローラ34が、ステップ54でリモートアダプタ31からコマ コードし、及びステップ56でそのコマンドを処理する。

[0032]

カスケードコマンド又は他のコマンドによって定義された動作が完了したとき デバイスコントローラ43は、リモートアダプタ31への転送のために、ステ ップ60で受取り(acknowledgemeni)を生成し、ステップ61、62及び63は ホスト24への転送のために、その受取りを通信リンクを逆に通してリモー アダプタ27及びホストアダプタ25に転送するための処理を表わす、

[0033]

なフィードバックを提供する。結果として、本発明を実施するための第1のリモ 今度はBCV/R1論理ボリューム34の動作を制御する第1のリモートサイト中のデ バイスコントローラ43へのコマンドを発行することができる。ステップ60か このように、この処理によって、ホスト24中で動作しているプログラムが、 ら63は、次に、ホストがすべてのそのような動作の完了に気づくように、 ートサイトで、図1のホスト33のようなホストに対する要件がなくなる。

[0034]

(第1のリモートサイト22)

0 特表20 冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

9 ~

വ

4

വ t

ケードコマンドセット中のコマンドを含む種々のコマンドを処理するためのモジ ための、多くのステータスレジスタ及びテーブル及びプログラムモジュールを含 む。コントロールモジュール70は、ESTABLISHコマンドを処理するためのESTモ ジュール71、DIFFERENTIAL SPLITコマンドを処理するためのDSモジュール72 及びREESTABLISHコマンドを処理するためのREESTモジュール73を含む、カス ュールを含む。ESTABLISH及びREESTABLISHコマンドは、カスケード動作から独立 図1をもう一度参照すると、第1のリモートサイト22は、本発明を実施す して、機能も実行する。

[0035]

の専用の部分を含む。R2論理ボリューム32の1つのそのような部分は、それぞ れる。それぞれのステータスワードは、トラックについての情報を含み、種々の 使用される。それらのビット位置の1つは、カスケード処理での使用のために割 論理ボリュームのそれぞれは、デバイスのための、個々の物理シリンダー、及 びそれらのシリンダー内のトラックのステータスについての情報を維持するため れのトラックについてのステータスワードを含む専用の部分74によって構成さ ステータスの機能について利用可能な多くのビット位置(bit position)のために り当てられるであろうし、またそのビット位置は、nがR2論理ボリューム32中 のトラックの数と等しい場合の1×nのテーブルである、プロテクションビッ テーブル(Protection Bit Table)75として複製されるであろう。

[0036]

トラックステータステーブル77も含む。トラックステータステーブル76及び ュームに割り当てられた4つの行を含む。BCV/RI論理ボリューム34がR2論理ボ リューム32に接続しているときの通常動作では、R2トラックステータステープ M2行中のビット位置は、R1論理ボリューム26中のトラックに対応し;及びM3行 ル76のM1行中のビット位置は、R2論理ボリューム32中のトラックに対応し; 第1のリモートサイト22は、R2トラックステータステーブル76及びBCV/R1 77のそれぞれは、複数の行、具体的には、4つのミラリングデバイス又はボリ 中のビット位置は、BCV動作モードで働いているBCV/RI論理ボリューム34中の トラックに対応する。同様に、BCV/R1トラックステータステーブル77は、

2003 06

冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

XI 論理ポリューム34のためのトラックステータスを識別するためのMI行中のビ ット位置及びR2論理ポリューム40のためのM2行中のビット位置を含む複数のミ ラー行を含む。

[0037]

PBビットレジスタ80は、カスケード動作の間に使用されることになるプロテ フラグ81は、システムがカスケードコマンドに応答して動作しているかどうか リモートサイト22設備中の異なる論理ポリュームの間でのデータのコピーを行 クションビットテーブル75中の選択されたビット位置を識別する。セッション を指示する。コピープログラム82は、後でより詳細に説明するように、

[0.038]

(動作-第1の動作モード)

の第1のデータ記憶設備からBCV/R1論理ボリューム34の形態の第2のデータ記 て、その動作モードの間にコピープログラム82がR2論理ボリューム32の形態 ションは、BCV/R1論理ボリューム34上のデータがR2論理ボリューム32上に記 第1のリモートサイト22は、リモートアダプタ31で受信された変更に従っ いのコンフィギュレー **筮されたデータと等しくなるような時間まで、同じ状態を保つ。その後、第1の 箘設備にデータをコピーする、そのような第1の動作モードで動作するように、** 最初にコンフィギュレーション設定が行われるであろう。 動作モードは、その同期を維持する。

[0039]

(動作-第2の動作モード)

そのような同期が達成された後はいつでも、ホスト24がDIFFERENTIAL SPLIT カスケードコマンドを発行するときに、第2の動作モードを確立することができ マンドを生成することも可能であろう。コマンドを発行することは、カスケード る。他には、第1のリモートサイトのオプションのホスト33は、 セッションを開始する。

応答として、デバイスコントローラ43は、BCV/RI論理ポリューム34をR2論

を転送することを可能にするであろう。それらの2つのボリュームが同期した後 **埋ボリューム32から分離し、図1のコピープログラムがBCV/R1論理ボリューム** 34からリモートアダプタ35を通じて、R2論理ポリューム40を含む図1の第 はいつでも、BCV/R1論理ボリューム34をリモートアダプタ35から分離し、そ 2のリモートサイト23によって表わされるようなデータ受信デバイスにデータ れをR2論理ポリューム32に再接続することによって、REESTABLISHコマンドは ンは、後続のDIFFERENTIAL SPLIT及びREESTABLISHコマンドを発行することによ システムを第1の動作モードに戻し、その後、コピープログラム82は、 論理ポリューム34をR2論理ポリューム32と再同期させるであろう。 って、CLOSE SESSIONコマンドが生成されるまで継続する。

$[0\ 0\ 4\ 1]$

コマンドをテストする。もし **何かエラーが存在するなら、その手頃は、技術上知られた報告手頃を使用して中** うに、リモートアダプタ27、通信リンク30及びリモートアダプタ31を通じ ホストアダプタ25が、ステップ90で異なるコマンドが受信されたことを認 **讃するとき、種々のあり得るエラーステータスを求めて、知られているように及** た転送のために、DIFFERENTIAL SPLITコマンドを転送するステップ92に制御が 図3は、より詳細にDIFFERENTIAL SPLITコマンドの受信に広答した動作を示す 止される。もしエラーが存在しないなら、制御部(control) 7 0 が応答できるよ び図2のステップ44に関して説明されるように、

[0042]

は、ロック(lock)が存在する間は、他のどのアプリケーションも選択された論理 ポリュームに影響を与えることを防ぐために、普通のロック動作(locking opera DIFFERENTIAL SPLIT(DS)モジュール72がステップ93で広答するとき、それ ードセッションが進行中であるかどうかを確認する。もしこれが、カスケード動 所定の位置に記憶されたデータに基づいてプロテクションビットテーブル(prote lion)を実行する。ステップ 9 4 は、セッションフラグ 8 1 をモニタし、カスケ 作の最初のディファレンシャルスプリット(differential split)動作であれば、 セッションフラグ81は、セッションは進行中でないことを指示する。制御は、

トリは、トラックに対応しており、それぞれのエントリは、任意の目的に割り当 てることができる複数のビット部分を含む。ステップ95は、開始されているセ うな有効な使用も識別するテーブル(図示せず)から、使用されていないピット 部分を選択する。そのようなビット位置を選択するための処理は、技術上周知で ステップ95は、メモリ中に維持されているそのようなファイルのコピーから 対応する情報を検索することができる。どのような形態でも、それぞれのエン ッションのための1つのビット位置を確立するため、異なるビット位置のどのよ ステップ ある。次に、DSモジュール72は、その選択されたビット位置をPB BITレジスタ データファイル(protection bits data file)を検索することができる。他には 例えば、 ction bits table)75を確立するステップ95に進行する。 95は、R2論理ボリューム32中の専用の部分74から、 80中に記録する。

[0043]

3 4に転送されることが防止される。しかし、配論理ボリューム32へのそれぞ てのBCV/RI論理ポリューム34のミラリング動作を終了させる。この実行によっ て、コピープログラム82が、R2論理ボリューム32からBCV/R1論理ボリューム れの転送は、R2トラックステータステーブル76のM3行中に記録され、それによ って、BCV/RI論理ボリューム34中に反映されなかったR2論理ボリューム26中 ステップ96は、普通のモジュールから基本的なステップを使用して普通のス プリット動作を実行するための手順を表わす。この動作の詳細は、図4に示され る。具体的には、ステップ101は、R2論理ボリューム32のためのミラーとし の変化が指示されるであろう。

[0044]

32からBCV/RI論理ボリューム34にコピーする。ステップ104は、次に、BC 953号中にの記載 **ートサイト22から第2のリモートサイトへのデータの転送がすべての識別情報** を含むように、どの識別テーブル(identification table)も、R2論埋ボリューム リ/R1論理ボリューム34を、第2のリモートサイト23のR2論理ボリューム40 のように、WRITE PENDINC動作を管理する。次のステップ103は、第1のリモ ステップ102は、前述の米国特許出願第08/842,

特表2002 冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

ဖ

~

വ

すなわち収論理ボリューム40にデータを転送することが可能 これによって、コピープログラム8 にデータをコピーするために割り当てる。 は、適当な宛先、

[0045]

リューム40に転送されることになる動作環境が確立される。ステップ105は トをセットすることによって、すべてのトラック中のデータが、図1のR2論理ポ どのような後続のDIFFERENTIAL SPLITの間でも、後述の他のパスが続くことが 全てのビッ 図4の動作が完了したとき、図3のステップ105は、BCV/R1トラックステ タステーブル77のM2ビット位置中のすべてのビットをセットする。 明確になるように、セッションフラグ81もセットする。

[0046]

ューム34からR2論理ポリューム40へのデータの転送を開始するコピープログ タステーブル77中の対応するM2ピットが消去される。全てのピットが消去され ステップ106は、ロックを解除する。ステップ107は、BCV/R1トラックス テータステーブル77中のデータ、特にM3行中の情報に従って、BCV/R1論理ボリ **転送が完了したホスト24中で動作する対応するアプリケーションプログラムを** ラム82を開始する。それぞれのコピー動作が起きると、BCV/R1トラックステー たとき、ステップ110は、完了ステータス(complete status)を通知(post)し 、対応するメッセージは、ホストアダプタ25に逆に転送され、それによって、 指示する。

[0047]

(動作-第1の動作モードに戻る)

ンドの受信を表わす。ステップ113は、もしエラーが検出されたなら処理を中 リモートアダプタ27及び通信リンク30を通じたリモートアダプタ31への すべてのデータがR2論理ボリューム40に転送された後はいつでも、システム することができる。図5のステップ112は、ホストアダプタ25でのそのコマ エラーが検出されなければ、ホストアダプタ25は、ステップ114を使用して オペレータは、別のカスケードコマンドを構成するREESTABLISHコマンドを発行 止する可能性がある、そのコマンドのコンテキスト及び文法のテストを表わす。

転送を制御するためのコマンドをデバイスコントローラ43に渡す

[0048]

ックステータステーブル76中の次に利用可能なデバイスのミラーの指示を有す ドの受信を指示する。ステップ116は、BCV/RI論理ボリューム34を、R2トラ グラムが更にデータを第2のリモートサイト23にコピーすることが防止される ことも、記憶されているデータに影響を与えない。ステップ120は、次に、BC //R1トラックステータステーブル77のためのM4ピット位置中のピットを有する ンドの間でシステムが第2のモードで動作しているとき、そのような変更はなさ タスレジスタテーブル76中のM3ビット位置の内容に従って、R2論理ポリューム それぞれの転送の間、コピープログラム82は、R1トラックステータスレジスタ デバイスコントローラ43は、ステップ115を使用して、そのようなコマン -ラ43は、DIFFERENTIAL SPLIT動作のための記憶設備として動作しているBCV/ BCV/R1論理ボリューム34(すなわち、M3ビット)に割り当てられたミラリング された位置にあったR2トラックステータステーブル76中のどのビットもマージ する。他のアプリケーションの間に、M4ビット位置は、BCV/R1デバイスのデータ になされたどのような変更も、通常記録する。しかし、DIFFERENTIAL SPLITコマ れない。続いてステップ120は、第1のリモートサイトがDIFFERENTIAL SPLIT **動作モードで動作していた間にR2論理ポリューム32になされたどのような変更** 女のステップ121は、ロックを解除し、またステップ122は、ホストアダプ テーブル76のM3ピット位置中の対応するピットを消去する。しかし、プロテク でホスト24に通知する。ここで、コピープログラム82が、R2トラックステー ローカルのBCVミラーとして追加する。ステップ117で、デバイスコントロ NI論理ボリューム34をNOT READY (NR)にセットする。これによって、COPYプロ 述のアプリケーションで説明したように、どのWRITE PENDING動作をセットする タ25に逆に転送される完了ステータス(complete status)を、ステップ123 も表わすR2トラックステータステーブル76のM3行中のビットを単に取得する。 32からBCV/R1論理ポリューム34にデータを転送するために利用可能になる。 ションビットテーブル75中の対応するピット位置は、変化しないままである。 どのWRITE PENDING動作も、ステップ118でINVALID状態にセットされる。

冗長データ配憶ユニットを通じてデータをカスケードにする方法および装置

[004.9]

(動作-R2論理ボリューム40への魯込み)

5に書き込む。ステップ134は、もしカスケード動作が進行中でないなら、バ 第1のリモートサイト22のデータ記憶設備が、BCV/R1論理ボリューム34か らR2論理ボリューム40にデータをコピーする第2の動作モードで動作する間隔 の間、ホスト24は、N1論理ポリューム26及びR2論理ポリューム32に書込み ム32に直ちに魯き込まれ、R2トラックステータステーブル76中の対応するエ ントリはステップ132で更新される。この場合、M3行中の対応するトラックビ ット位置は、WRITE動作がR2論理ポリューム32への転送を引き起こしたが、BCV に更新されるであろう。次に、システムは、セッションフラグ81を見て、カス ケード動作が進行中であったかどうかを確認するであろう。もしそうであったな ら、ステップ133は、制御をステップ134に移し、それによって、プロテク ョンピットテーブル7 5 及びトラックステータステーブル7 6 中のM3行が同一 の情報を有するように、対応するステータスをプロテクションピットテーブル7 イパスされる。ステップ1'35は、次に、その処理によって制御部70が魯込み 動作を発行し続けることができる。リモートアダプタ3.1は、図6のステップ1 変更されたデータをBCV/R1論理ポリューム34に転送することによって完了す 3 0 でWITEデータを受信する。その情報は、ステップ131でR2論理ボリュー /RI論理ボリューム34中の対応する変更を引き起こさなかったことを示すよう 動作を完了させる、そのような処理を表わす。通常動作の場合は、RRITE動作は もし、BCV/RI論理ボリューム34が接続されていなければ、WRITEリクエス WRITE PENDINGリクエストにされるであろう なれば

[0050]

(動作ー連続ディファレンシャルスプリット動作)

PLITコマンドを発行する。デバイスコントローラ43がそのコマンドを受信する 図4のステップ101に関して説明したものと同様の方法で、ステップ94がス ができる、いくつかの迫加の時間間隔の後、ホスト21は、他のDIFFERENTIAL システムのオペレータが定義することができ、分、時間又は日で測定するこ それの制御は、図3に示された手順に再び移される。しかし、この場合、

特表2002 冗長データ記憶ユニットを通じてデータをカスケートにする方法および装置

ようにされる。ステップ142は、ステップ102と同様に、全てのWRITE PEND 転送が目的の、コピープログラム82のどのような更なるアクションもできない テップ141に分岐するようにセッションフラグ81はセットされ、R2論理ボリ これによって、R2論理ボリューム32からBCV論理ボリューム34へのデータの ューム32へのミラーとしてのBCV/RI論理ボリューム34の動作を終了させる。 ING動作を管理する。

[0051]

は、処理されている最後のDIFFERENTIAL SPLITリクエストに続いて変更されたR2 論理ポリューム32及 次に、DSモジュール72は、ステップ143を使用して、プロテクションビッ トテーブル75を、BCV/R1トラックステータステーブル77中のM2行にコピーす る。次に、108モジュール72は、プロテクションビットテーブル75を消去する これらの動作の後、BCV/R1トラックステータステーブル77中のM2ビット位置 論理ボリューム32中のそれらのトラックを指示する。プロテクションビットテ ットテーブル75中の適当なピットをセットすることによって適切に記録される び34のロックを外すと同時に、どのWRITE PENDING動作も、プロテクションビ ーブル75中のプロテクションビットを消去することは、 ことを確実にする。

[0052]

トラックの中のデータのみが送信されるであろう。すなわち、セッション中の最 タに対応する、増加する量の情報のみを送信するということは、このコマンドの 常かなり小さいであろう。例えば、もしそのような間隔の間に、R2論理ポリュー ム32の同じトラック上の異なる領域への、10回の書込み動作があったとした にシフトし、その結果、コピープログラム82が、BCV/RI論理ボリューム34か 性質である。更に、通信リンク36を通じて実際に転送されたデータのトラック DSモジュール72中の次の制御は、ステップ93でかけられたロックを解除す るためにステップ143からステップ106にシフトし、そしてステップ107 らR2論理ボリューム40にデータを転送できるようになる。しかし、変更された 初の動作の後の全てのDIFFERENTIAL SPLIT動作は、変更されたトラック中のデー の数は、連続したDIFFERENTIAL SPLIT動作の間に起こる書込み動作の数より、

542 特表2002-冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

ဖ ~

Ŋ

クに発生するとき、通信リンク36のためのバンド幅の要件を劇的に減少させる ら、BCV/RIデバイス34は、最も新しい情報のみを含み、そして1つのトラック 書込み動作のみが、第2のリモートサイト23中のR2論理ポリューム40を更新 するために必要になるであろう。そのような反復する書込み動作が単一のトラッ として使用することができるようなレベルに更にバンド幅の要件を減少させるこ 股的にわかる。多くのアプリケーションでは、インターネットを通信リンク3 ことができる。要件を十分に減らすことができ、そして、TIラインのような、 なくとも次のより低いレベルの通信バンド幅を使用することができることは、 とができる。

[0053]

のリポジトリとして、第2のリモートサイト23を使用することが望まれる限り 冗長なコピーを維持する必要がなくな DIFFERENTIAL SPLIT及びREESTABLISHコマンドの順番は、冗長なコピーのため のカスケードコマンドは、レジスタ 8 0 中のPBビット位置の割り当てを開放し、 ったら、ホストは、所定のCLOSE SESSIONコマンドを生成することができる。 カスケード動作を完了させるためにセッションフラグ81を消去する。 mj 、反復して継続することができる。

[0054]

本発明は、ローカルのプロダクションサイト21が、プロダクションデータ記 **懲設備中でデータを周期的に変更するための装置を表わすような好適な実施の形** ションサイトでデータをミラリングするための第1のデータ記憶装置を含む。BC V/R1論理ボリューム34の形態の第2のデータ記憶装置は、2つの動作モードを 有する。1つでは、BCV/R1論理ボリューム34は、R2論理ボリューム32のため のミラーとして働く。DIFFERENTIAL SPLIT又は同等のコマンドを発行することに よって開始される他のモードでは、BCV/R1論理ボリューム34は、R2論理ボリュ ーム32から分離され、BCV/R1トラックステータステーブル77中で識別された データをR2論理ボリューム40、又はBCV/R1論理ボリューム34からのデータの 態に関して説明してきた。本発明によると、第1のリモートサイトは、プロダク ためのレシーバとして働く第2のリモートサイト23中の他の手段に伝達する。

[0055]

プロテクションビットテーブル75によって表わされるデータ変更記録装置は ム40に転送されている間に、R2論理ボリューム32へのどのような変更も記録 する。制御がBCV/RI論理ボリューム34を逆に第1の動作モードにシフトしてい るとき、データの変更の記録は、それらのデータブロック又はR2論理ボリューム クを識別する。この変更されたデータの記録は、システムが再びBCV/R1論理ポリ ム34から第2のリモートサイト23のようなデータ受信サイトに転送されるよ 3 2からBCV/R1論理ボリューム34に転送しなければならない他のデータブロッ 第2の動作モードの間中、BCV/R1論理ボリューム34の内容がR2論理ボリュー ューム34の動作を第2の動作モードにシフトさせるとき、BCV/R1論理ポリュー うなトラックを更に制御する。

[0056]

図1から6に示された具体的に開示された実施形態に行うことができる多くの る、他の順番でもよいであろう。そのような順番は、動作モードの間のそれぞ れの転送をシステムのオペレータが開始するために、マニュアルで制御すること 変形がある。図1は、例えば、テーブル及びモジュールの特定の編成を表現する 第1のリモートサイト22の全体の動作から減じることなく、その情報の異な 論理ボリュームに関連する動作を更に表現する。一般的に、ネットワークは、図 1から6に示される動作及び装置を繰り返すことによって、それぞれをこれと同 じ方法で扱うことができる複数の論理ポリュームを含むであろう。更に、種々の モジュールの説明は、説明の目的のために特定の順番を確立するが;他の補助的 な動作のために必要なような他の動作のいくつかの削除又は追加も含むことがで 他には、カスケードセッションのすべて又は部分を自動制御とするこ 本発明の真の精神及び範囲内の、すべての変形及び変更を包含することが、添付 ともできよう。このように、本発明は特定の実施形態に関して開示してきたが、 るコンフィギュレーションで置きかえることができよう。図1から6は、 の請求項の目的である。 ができる。

[図面の簡単な説明]

[図1]

プロダクション設備及び2つの地理的に遠隔の設備を含むデータ処理ネットワ

冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

ークのプロック図である。

[図2]

プロダクション設備からリモート記憶設備へのコマンドの転送を示すフロー図

である。

[<u>8</u>3]

DIFFERENTIAL SPLITコマンドに応答した図1のネットワークの動作を示す図で

88

[図4]

図3で示された動作で使用される手順の詳細を示す図である。

[図2]

RE-ESTABLISHコマンドに応答した図1のネットワークの動作を示す図である。

[9図]

DIFFERENTIAL SPLITコマンドを処理した後に起きる他の動作を示す図である。

53

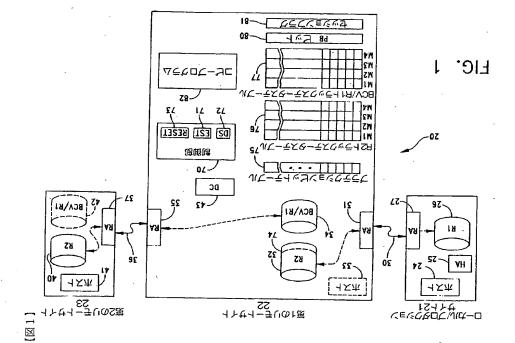
526

特表2002-542

冗長データ記憶ユニットを通じてデータをカス ケードにする方法および装置

9

5 2 42



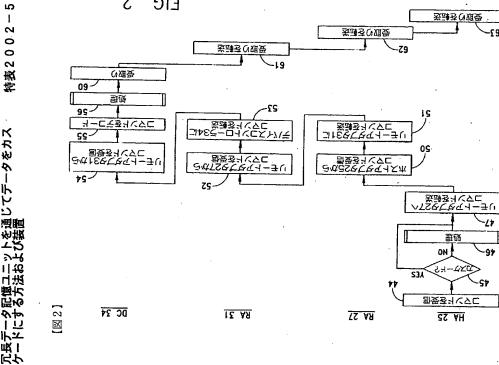
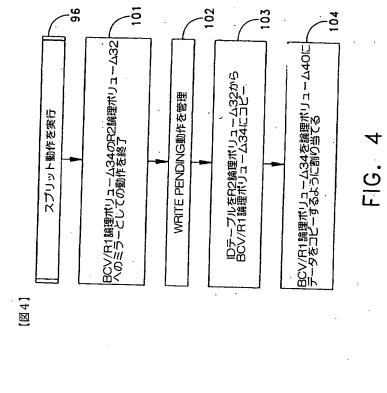


FIG.

7

9



ディファレンシャルスプリットリクエス トを発見:ロックをかける

ディファレンシャルスブリットリクエストを発行が離

2

비

カスケードセッションは低行中で

YES

プロテクションピットテーブル75を確立

RZ稿理ポリューム32へのミラーとして 808CV/R1橋理ポリューム31の動作を 87:BCV/R1橋理ポリューム34部2篇 理ポリューム40ペデータをコピーするだ めに割り当てる

195

2

100

105/

ロックを解除

-BCV/R1トラックステータステーブルT7 に従ってBCV/R1線壁ボリューム34から R2誘電ボリューム40ヘデータをコピー

- 協理ボリューム34からR2へ完了ステータスを通知

完了ステータスを発見し、再接続

M

F1G.

ステータステーブル77 トをセット:セッション

テクションピットテーブル75をBCV トラックステータステーブルイ7に デブロテクションピットテーブル75

WRITE PENDING動作を管理

スプリット動作を実行

-142

特表2002-542526

冗長データ配億ユニットを通じてデータをカス ケードにする方法および装置

ホスト24からディファレン シャルスプリットコマンドを 受信

HA

[83]

特表2002-542526

冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

[図5]

526

-542

特表2002

中中

FIG. 5

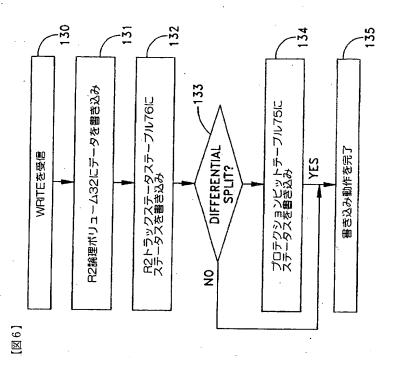
完了ステータスを発見し、 再接機

ータスを通知

完了ステ

121

ロックを解除



G. 6

冗長データ配億ユニットを通じてデータをカスケードにする方法および装置

[提出日] 平成13年3月7日(2001.3.7)

[手続補正1]

[補正対象書類名] 明細書

【補正対象項目名】特許請求の範囲

[補正方法] 変更.

[補正の内容]

[特許請求の範囲]

(請求項1] データ処理システム中で、データを変更する手段と、を受信する手段と、を有するデータネットワークであって、

データ

- A)データを受信するために前記データを変更する手段へ接続するための第1のデータ記憶装置と、
-) 第2のデータ記憶装置と、
- C)前記第1のデータ記憶装置中で前記データを変更する手段が行う変更の位置を識別するための第1及び第2の変更記録手段と、
- D)前記第1の変更記録手段中に記録された変更に従って、その間に前記第2のデータ記憶装置が前記第1のデータ記憶装置からデータを受信するような第1の動作モードを確立するための第1の動作手段と、
- E) 第2の動作モードが確立されているときに前記第1のデータ記憶装置中の記録された変更を反映している前記第2の変更記録手段中に記録された変更に従って、前記第2のデータ記憶装置から前記受信する手段にデータをコピーするための前記第2の動作モードを確立するための第2の動作手段と、を有することを特徴とするデータネットワーク。

(請求項2] 前記第1及び第2の変更記録手段は、それぞれ第1及び第2のテーブルを含み、前記データネットワーク内の転送は、データブロックとして起こり、前記第1のデータ記憶装置は、それぞれのデータブロックについての別の情報のための記憶装置を含み、前記第2の変更記録手段中の前記第2のテーブルは、当該別の情報を受信し、前記第2の動作手段は、前記第2のデーブルの内容を、前記第2のデーグルの内容を、前記第2のデータ記憶装置から前記データを受信する手段へのデータの転

送を制御する第3のテーブルに転送することを特徴とする間求項1に記載のデータネットワーク。

ノボンドン ブ。 【静水項3】 前記データ記憶設備内の転送は、データブロックとして起こ

り、前記第1のデータ記憶装置の一部は、それぞれのデータブロックへの変更に ついての情報を含み、前記データ記憶設備は、第3の変更記録手段を更に含み、 及び、前記第2の動作手段は、第2のモードが確立されているときに、前記第1 のデータ記憶装置の一部の内容を前記第3の変更記録手段に転送することを特徴 とする請求項1に記載のデータネットワーク。 [請求項4] その間に制御手段が交互にかつ反復して前記第1及び第2の動作手段を相互に排他的に使用可能にするようなカスケードセッションを実行するための、前記変更する手段から反復して生成される第1及び第2のコマンドに応答する前記制御手段を更に含むことを特徴とする請求項3に記載のデータネットワーク。

【請求項6】 データを変更する設備と、データを受信する設備の間でデータを転送するための方法であって、

- A)変更されたデータを受信するための第1のデータ記憶装置を確立するステップと、
- B)第2のデータ記憶装置を確立するステップと、
- C) 第1及び第2の位置に、データを変更するための設備によって作り出される 前記第1のデータ記憶装置中の変更の識別を記録するステップと、
-)) 前記第1の位置中の識別された変更に従って、その間に前記第2のデータ記

38

憶装置が前記第1のデータ記憶装置からデータを受信するような第1の動作モー ドを確立するステップと、 E) 第2の動作モードが確立されているときに識別される第2の位置中の変更に 従って、前記第2のデータ記憶装置から前記受信する設備にデータをコピーする ための第2の動作モードを確立するステップと、を有することを特徴とする方法 データ転送は、データプロックとして起こり、前記第1のデ 夕記憶装置は、それぞれのデータブロックについての別の情報のための記憶装 置を含み、変更を識別する前記ステップは、変更の識別を記録するための第3の 位置を確立することを含み、前記第2の動作モードの確立は、前記受信する設備 へのコピーができるようにするために、第2の位置の内容を第3の位置へ転送す ることを特徴とする請求項6に記載の方法。 [請求項7]

データ転送は、データブロックとして起こり、前記第1のデ ータ記憶装置の一部は、それぞれのデータブロックへの変更についての情報を含 み、前記第1のデータ記憶装置の一部の内容は、第2の動作モードが確立されて いるときに、最初の識別情報を提供するために前記第2の位置に転送されること を特徴とする請求項6に記載の方法。 [請求項8]

[請求項9] 前記変更する設備からの第1及び第2のコマンドの反復した 生成は、その間に前記第1及び第2の動作モードが相互に排他的に使用可能にさ れるようなカスケードセッションを反復して実行することを特徴とする請求項8 に記載の方法。

ドの最初の動作は、当該コピーが前記第2のデータ記憶装置中のすべてのデータ を前記受信する設備に転送するようなカスケードセッションにおける最初の反復 の間に、前記第2の位置を初期化し、及び、第3の位置中のすべてのデータを第 データは、コピーによって転送され、前記第2の動作モー 1の状態に設定することを特徴とする請求項9に記載の方法 [請求項10]

[国際調査報告]

冗長データ記憶ユニットを通じてデータをカスケードにする方法および装置

\$ 99/23	
PCT/U	
	PCT/US 99/23

	INTERNATIONAL SEARCH REPORT PET/	ntern at Application No PCT/US 99/23939
PP-341	L CLYBERONION OF BITMET WATTER	
According (According to International Patent CheesBookin (IPQ) or to both natural clearification and IPO Per part na seast-Geren	
IPC 7	kitherin des any estem nevertee (desattemen system televest by desattemen system) 100.7 ± 0.06 F	,
Documents	Doornentiden eersted oles fest mitheren doornentiden to be askeri bis esh doornerte are broaded. In the false seembed	so fields asserted
Bectroto	Describ sha hees consided daing the Membrais esseth forms of data hase sea, when provided, essech horn wed	orre lawed)
100	A PACHIBERT KYMEN GSER YWAR	
Company.	Ohalen of document, with indication, whose appropriate, of the relevant presenges	Pladewart to dishs He.
ec .	WD 97 45790 A (EMC CORPORATION) 4 December 1997 (1997-12-04) page 54, line 15 -page 55, line 15; figure 22	1,8,20, 27
- <	WO 98 20419 A (VINCA CORPORATION) 14 May 1998 (1998-05-14) page 5, line 15 -page 7, line 34	1,8,20,
æ	US 5 799 147 A (SHAWHON) 25 August 1998 (1998-08-25) the whole document	1,8,20,
«	US 5 857 208 A (OFEK) 5 January 1999 (1993-01-05) abstract	27, 20,
	<u>'</u>	-
×	Pusher documents are blend in the construction of box G.	ase Method in graceat,
Parter .	* Spacks exegorise of ched documents :	serthe International Strip date

X Petert brilly members are feind in meas.	The last countried published of the Bar Newsberrid Birty dis- bards to understand by principle as the rest first plan. In the countried published as the rest of the bards of the producer of a published by the consideration of the countried to considerate the bard to considerate the rest of the countried by the countried to be countried to be considered to be considered to be countried to be countried to be considered to the considered to be considered to be considered to the	Date of milling of the interrediental remoch report 30/03/2000	Authoriton other Absalom, R
X Pusher documents are Literal in the confination of best C.	* Special empeyane of devel decorrance; **Constraint the original proposal acts of which is not constraint to the original proposal acts of the original proposal acts of the original proposal acts of the original proposal propo	Date of the extual correlation of the International seasoh 23 March 2000	Nieses and melling dischess of the St. P. B. Bill Phaintiese; 2 B. Capper Phant Clear, P. B. Bill Phaintiese; 3 H. L. Capper (1984) P. S. Bill Phaintiese; 1 Fax (491-10) 340-2016

page 1 of 2

page 2 of 2

冗長データ記憶ユニットを通じてデータをカス ケードにする方法および装置

特表2002-542526

6	
REPORT	
TIONAL SEARCH R	
INTERNATIONAL	
INTERN	

	16-01-1998 24-03-1999 10-11-1998 29-05-1998	11-01-1998 12-12-1998 12-12-1998 12-12-1998 11-08-1997 14-05-1988 15-01-1994 14-12-1986 14-12-1986
05-01-1998 24-03-1999	10-11-1998 29-05-1998 22-12-1998	22-12-1998 29-05-1908 22-12-1998 11-08-1987 20-08-1986 05-01-1994 06-11-1994 06-11-1994 11-09-1986
10-11-1998 29-05-1998	22-12-1998	22-12-1998 11-08-1987 24-05-1998 05-01-1994 04-12-1986
29-05-		22 22 11 68 69 69 69 69 69 69 69 69 69 69 69 69 69
5835953 A 5174098 A 5852713 A		20120 20120 246 747 847 847 847 847 847 847 847 847 847
5852713		466620 1237201 0191036 6001447 61502846 9308684
1.	씵ㅣ	25655
2 25 2		事 コウェッショ
88 8		985
04-12-1997 14-05-1998 25-08-1998		13-02-1986
	_	
(4 4	•	< <
9745790 98 <u>2</u> 0419 5799147		US 5857208 UO 8601018
2 3 5	÷	{ 8

フロントページの続き

ドターム(松岑) 5B018 GA04 HA05 MA12 QA15 5B082 DE03 GB02 5B083 AAOS CC04 CD11 EE08

JEST AVAILABLE COPY